Brodogradnja Volume 75, Number 2 (2024) 75203

journal homepage: www.brodogradnja.fsb.hr

Brodogradnja

An International Journal of Naval Architecture and Ocean Engineering for Research and Development

Ship collision avoidance decision-making research in coastal waters considering uncertainty of target ships

*Jianjie Gao, Yuquan Zhang**

1 Navigation College, Jimei University, Xiamen 361021, China
2 Hebei Jiaotong Vocational and Technical College, China

ARTICLE INFO

Editor-in-Chief: Prof. Nastia Degiuli
Associate Editor: PhD Ivana Martić

Keywords:
Ship collision avoidance
Automatic identification system data
Trajectory clustering
Ship maneuverability
Trajectory prediction

ABSTRACT

Ship collision avoidance has always been a concern and it is crucial for achieving safe navigation of ships at sea. There are many studies on ship collision avoidance in open water, but less attention on coastal waters considering the uncertainty of target ships due to the complexity of the environment and traffic flow. In this paper, collision avoidance decision-making research in coastal waters considering the uncertainty of target ships was proposed. Firstly, accurate ship trajectories are obtained by preprocessing the raw Automatic Identification System (AIS) data. Subsequently, the processed trajectories are clustered using the Ordering Points to Identify the Clustering Structure (OPTICS) algorithm and Hausdorff distance, acquiring a dataset for trajectory prediction of target ships. Then, a mixed Gaussian model is utilized to calculate the prior probability distribution of the prediction model, thus establishing a trajectory prediction model that considers the uncertainty of the target ship. Finally, ship maneuverability is simulated using the Mathematical Model Group (MMG) and Proportion Integration Differentiation (PID) models, and a collision avoidance decision-making model for ships is constructed. The proposed algorithm has been tested and verified in a case study. The results show that the approach effectively predicts the trajectory of the target ship and facilitates informed collision avoidance decision-making.

1. Introduction

With the rapid development of global economic integration, the role of maritime trade in global trade is becoming increasingly prominent, becoming an important component of the comprehensive transportation system, and its position cannot be replaced by other modes of transportation [1]. At the same time, with the rapid development of science and technology, ships are moving towards intelligence, large-scale, and high-speed development, and the maritime navigation environment has become increasingly complex. However, in recent years, the incidence of water traffic accidents has remained high in global waters [2]. The occurrence of maritime accidents can cause significant economic losses, casualties, and environmental pollution [3]. In

* Corresponding author.
E-mail address: zyqjytj@126.com

http://dx.doi.org/10.21278/brod75203
Submitted 09.09.2023.; Accepted 08.01.2024.
ISSN 0007-215X; eISSN 1845-5859

January 2018, the Panamanian oil tanker "SANCHI" collided with the "CF CRYSTAI" cargo ship at the mouth of the Yangtze River, causing serious damage to the "CF CRYSTAI" ship, resulting in three deaths and 29 missing people. In December 2020, the container ship "Xinqisheng 69" collided with the Papua ship "OCEANA" at the mouth of the Yangtze River, resulting in 5 deaths. The International Maritime Organization (IMO) believes that almost all maritime safety accidents are directly or indirectly caused by driver errors, i.e. human factors [4-5]. Therefore, it is a consensus in the maritime industry that human factors are the main cause of collision accidents. Improving the level of intelligence and automation during ship navigation can effectively reduce the impact of human factors, so ship intelligence is considered an important means to reduce collision accidents. ship intelligence refers to a ship with intelligent functions, specifically including the use of artificial intelligence, communication technology, and other technological means to make the ship self-control, autonomous navigation, and self-operation intelligent.

Intelligent navigation is one of the core functions of intelligent ships, and collision avoidance decision-making in intelligent navigation is a crucial technology in the development process of intelligent navigation. At present, the main methods used for ship collision avoidance include geometric analysis, heuristic methods, and artificial intelligence methods. However, most studies focus on path planning or proposing collision avoidance decision plans. The research on collision avoidance decision-making methods considering the uncertainty of target ships based on true AIS data is a very important topic. This article proposes a collision avoidance decision-making research in coastal waters considering the uncertainty of target ships, starting from the process of AIS data preprocessing, trajectory clustering, trajectory prediction, and ship collision avoidance decision-making.

The contents of the paper are arranged as follows: A detailed literature review of trajectory clustering, trajectory prediction, and collision avoidance is presented in Section 2. The proposed collision avoidance decision-making method is presented in Section 3. The case study and analysis are carried out in Section 4. Finally, the conclusion is drawn in Section 5.

2. Literature review

2.1 Trajectory clustering

The clustering of ship trajectories based on AIS data is the foundation of ship collision avoidance decision-making. Trajectory clustering can divide highly similar trajectories into the same trajectory cluster by determining the degree of similarity between trajectories, thus achieving ship route modeling [6]. From the perspective of clustering algorithm principles, clustering algorithms can be mainly divided into Prototype-based clustering algorithm, Density based clustering algorithm, Hierarchical based clustering algorithm, and Map based clustering algorithm.

The basic idea of the prototype-based method is to randomly select some trajectories as clustering prototypes. By calculating the distance between each trajectory and each clustering prototype, the trajectories are assigned to the nearest prototype, and then the position of the prototype trajectories is updated to gradually improve clustering quality. The typical algorithms include K-means and K-medoids. Zhao et al. [7] used a dynamic time warping algorithm to measure similarity values and successfully clustered AIS data in the Zhoushan water area based on the k-medoids clustering algorithm. Zhen et al. [8] improved the Hausdorff distance and used the k-medoids algorithm to cluster ship trajectories, taking into account the heading information of trajectory points. Bai et al. [9] proposed a ship route design method based on the K-means and ant colony algorithm, which uses the K-means algorithm to cluster routing in underwater.

The classical algorithms based on density clustering include DBSCAN, OPTICS, etc. Palotta et al. [10] proposed a density clustering algorithm DBSCAN for ship route extraction and anomaly detection, which is used for the perception of ship behavior trends at sea. Yang et al. [11] proposed a density-based DBTCAN trajectory clustering algorithm with noise that can directly cluster similar ship trajectories to address the issues of low efficiency, low detection accuracy, and loss of local features of ship trajectories clustering algorithms. Zhao et al. [12] proposed a method based on Douglas Pucker (DP) compression and improved adaptive DBSCAN to improve the clustering performance of ship trajectory data characterized by large data volume and distribution complexity, which demonstrated excellent performance in both time and quality. Shen et al. [13] addressed the issue of traditional clustering methods being unable to distinguish trajectories in opposite directions. They used cosine Hausdorff distance to construct a similarity function, adopted an adaptive compression ratio-based trajectory simplification method to simplify trajectories, and finally clustered trajectories using the DBSCAN algorithm.

Based on hierarchical clustering algorithms, it can be divided into top-down aggregation clustering (AC) and bottom-up divisive clustering (DC). Cao et al. [14] introduced Frechet distance to measure the similarity between AIS trajectories and used principal component analysis to decompose the generated similarity matrix to determine the number of clusters in the final clustering level. Finally, the practicality and effectiveness of the improved hierarchical clustering algorithm were verified using the AIS dataset in the Yangtze River region as an example. Vicente et al. [15] used a hierarchical clustering algorithm to cluster the trajectory information of European cruise ships to understand their traffic behavior characteristics. The clustering results matched the information published by the European Cruise Association. Wang et al. [16] developed a fast clustering model for ship trajectories based on hierarchical modeling. Each ship state establishes its own trajectory similarity model and recursively clusters ship trajectories from top to bottom, avoiding the drawbacks of existing ship clustering models such as cumbersome calculation, high time complexity, and difficult parameter adjustment process. Zhao et al. [17] proposed an adaptive hierarchical clustering method based on the DBSCAN algorithm. Experimental results have shown that this method can better identify ship trajectories.

In these studies, clustering algorithms can mine additional navigation environment information from a data perspective. However, these algorithms have the disadvantage of requiring a declaration of the number of clusters, being sensitive to initial state and noise data, and being unable to cluster irregularly shaped targets. The density-based clustering algorithm is very sensitive to parameter settings, and the hierarchical clustering algorithm has high computational complexity and is time-consuming. For different measurement and aggregation methods, the clustering results may have significant differences [18].

2.2 Trajectory prediction

Constructing accurate and reliable ship trajectory prediction models can serve as the basis for ship collision avoidance decisions and ensure the safety of ship navigation [19]. Ship trajectory prediction methods can be divided into three categories: methods using physical models, methods using learning, and methods using planning.

In terms of ship trajectory prediction models based on physical models, Xiao et al. [20] considered the accuracy, efficiency, and practicality of prediction methods, adopted particle filtering to track ship motion patterns, and combined them with physical motion models to calculate ship motion directions and ultimately output ship motion positions, thereby achieving ship trajectory prediction. Perera et al. [21] proposed a ship maneuvering model based on curve motion, taking into account both the position trajectory, velocity, and acceleration of the ship, and further designed an adaptive Kalman filter for ship trajectory prediction using Gaussian white noise.

Learning-based methods for predicting ship trajectories are theoretically based on supervised or unsupervised learning, which continuously learns the temporal characteristics of ship motion trajectories from AIS data to fit the predicted trajectories. Virjonen et al. [22] applied the k-nearest Neighbors (KNN) algorithm to ship trajectory prediction. This study used the current ship's navigation trajectory sequence to search for the nearest K ship trajectory points, and calculated feature similarity for the corresponding trajectory sequence of K points. The highest similarity was used as the final prediction result. Liu et al. [23] first clustered historical ship trajectories, and then independently modeled the X and Y directions of the trajectories using the Gaussian process regression method to achieve ship trajectory prediction. Rong et al. [24] proposed a probabilistic trajectory regression prediction model based on the Gaussian process model.

In terms of planning-based ship trajectory prediction models, Lazarowska [25] designed a ship path planning algorithm based on discrete artificial potential fields for autonomous ships, which can ensure real-time generation of safe paths in environments with static and dynamic obstacles. Silveira et al. [26] first constructed a graph structure based on AIS data, and then solved the ship safety path using the Dijkstra shortest path algorithm. Karbowska [27] proposed a ship trajectory planning generation method based on a beam search strategy, which can generate a trajectory tree composed of the ship's avoidance strategy. Through the final experiment, it was found that this method can find shorter trajectory paths.

In summary, in the short-term prediction research of ship trajectories, a common solution is to better predict the future motion status of ships by integrating traffic context features. However, methods using physical models have significant difficulties in encoding traffic environment features, which cannot better capture the environmental and interactive factors of ship motion. In predicting ship trajectories over a relatively medium to long-term time span, the above methods are difficult to ensure prediction accuracy, and most ship trajectory prediction studies are focused on open sea areas, with less research on predicting ship trajectories within busy port waters.

2.3 Collision avoidance

Ship avoidance decision-making is one of the key technologies in intelligent navigation, aiming to plan a path that can successfully avoid all incoming ships based on specific avoidance principles [28]. Ship collision avoidance methods are divided into two types of collision avoidance decision-making methods: global path planning and local path planning. Global path planning refers to the comprehensive planning of a ship's navigable path based on multiple factors such as the natural environment of navigation, water vessel traffic situation, and the ship's navigation plan. It searches for a collision free optimal path from the starting point to the endpoint based on set evaluation criteria. Local collision avoidance decision-making refers to providing a collision avoidance trajectory for a ship based on considering the local dynamic and static environment, to ensure that the ship has good obstacle avoidance ability.

In terms of global path planning, researchers have conducted extensive research. The global path optimization mainly includes the virtual potential field method, graph search algorithm, and group optimization algorithm. Huang et al. [29] proposed a variable APF method to achieve path planning, collision avoidance, and formation of unmanned aerial vehicles in the area of concern, taking into account the coupling problem between unmanned aerial vehicle dynamics and path planning strategies. Wang et al. [30] used the APF method to design a method for autonomous underwater vehicles to avoid collisions with the seabed and obstacles, but there is still a local minimum problem, and dynamic obstacle avoidance is not considered. The A * algorithm is representative of the field of USV global path optimization. Singh et al. [31] proposed the A * algorithm that considers safe distance and applied it to sea environments with complex ocean current interference. By reducing the range of searchable sea surfaces, the computational efficiency is further improved. In terms of swarm optimization algorithms, Xue et al. [32] combined particle swarm optimization with a sine cosine algorithm to design a quasi-reflection based on the sine cosine particle swarm optimization algorithm. Xie et al. [33] proposed an underactuated surface ship prediction collision avoidance method based on an improved Tendon whisker search algorithm, which considers the constraints of ship motion models and collision avoidance rules.

Domestic and foreign scholars commonly use methods for local path planning, including geometric analysis, heuristic methods, and artificial intelligence methods. The geometric analysis method determines the responsibility and behavior of vessel avoidance through quantitative analysis of the "Collision Avoidance Rules". Through quantitative analysis and evaluation of information such as vessel position, speed, and heading feasible avoidance intervals are determined, and feasible avoidance decisions are obtained through random or iterative methods [34]. Liu et al. [35] also defined ship avoidance actions based on the "Collision Avoidance Rules", and finally solved feasible avoidance decisions with the goal of not infringing on the ship domain; Wang et al. [36] divided the multi-vessel encounter situation into multiple groups and quantitatively analyzed the encounter situation in the "Collision Avoidance Rules" based on the relative speed and navigation of the two vessels, to study the avoidance decision-making under the multi vessel encounter situation, and to obtain the possible avoidance behavior of each vessel, laying the foundation for the multi vessel encounter. The heuristic method uses the "Collision Avoidance Rules" as constraints to obtain a feasible avoidance space, and combines safety, economy, and other factors to construct an objective function. Through the heuristic function, the search is guided to obtain the best avoidance decision in the solution space. Fiskin et al. [37]

defined the types of encounter situations and ship avoidance responsibilities based on the COLREGs, and used meta-heuristic optimization algorithms for decision-making optimization after determining the timing and behavior of avoidance. Xie et al. [38] regarded the "Collision Avoidance Rules" as a constraint condition for collision avoidance, combined with ship motion and control models to construct an objective function, and used the Taurus virtual optimization algorithm to find the optimal decision. Lazarowska [39] also determines the optimization space for collision avoidance decisions based on the "Collision Avoidance Rules", and then uses the ant colony algorithm for decision optimization. Common methods of artificial intelligence include deep learning and reinforcement learning. Chen et al. [40] proposed a path planning and maneuvering method based on Q-learning. Firstly, a ship maneuvering motion model was established using the Nomoto model, and then distance, obstacles, and other factors were used as rewards and punishments. Q-learning was introduced to learn the action reward model for judging ship autonomous collision avoidance decisions. Gao et al. [41] proposed an improved deep Q-Learning algorithm, which is used to solve the limitations of the DQN algorithm in adaptability and low success rate in ship path planning.

3. Methodology

3.1 Methodological overview

In this research, a decision-making research in coastal waters considering the uncertainty of the target ship is proposed. The overall flowchart of the research is illustrated in Fig. 1. The research consists of four main parts: data preprocessing, trajectory clustering, trajectory prediction, and collision avoidance.

There are plenty of outliers and missing data in raw AIS data, due to the AIS messages being easily affected by bad weather, blocked communication channels, equipment errors, etc. To obtain higher-quality data, it is essential to perform preprocessing steps such as decoding, cleaning, and interpolation on the AIS data. In this paper, the OPTICS clustering algorithm is used to cluster ship trajectories after the data preprocessing. The aim is to classify ship trajectories with different navigation behaviors, to provide a more focused dataset for trajectory prediction that accounts for specific navigation behaviors. The accuracy of trajectory prediction will improve after clustering. The future trajectory of the target ship is particularly important for ship collision avoidance research. In this study, the trajectory prediction based on clustered ship trajectories is implemented using the Gaussian Mixture Model (GMM), providing a reference for ship collision avoidance research. After obtaining the future trajectory of the target ship, this paper will establish the objective function of collision avoidance. The optimal decision-making for avoidance is determined by solving the objective function while considering ship maneuverability.

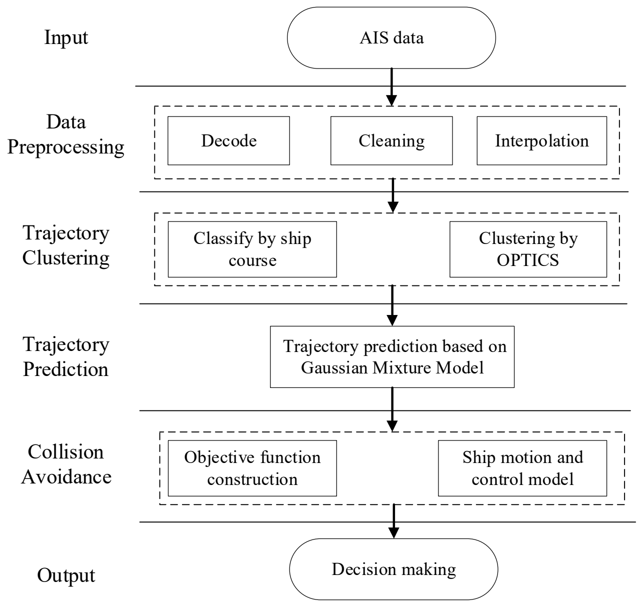


Fig. 1 The flowchart of the collision avoidance model

3.2 Data preprocessing

To improve the data quality and reduce the impact of noise on the results, it is necessary to preprocess the AIS data. The AIS messages are composed of compressed ASCII codes to facilitate transmission. The necessary data can be extracted after the decoding of the AIS messages. The AIS consists of 27 different types of message packets, each serving different purposes. In this paper, the focus is on decoding the information of Maritime Mobile Service Identity (MMSI), longitude, latitude, course, and speed.

The AIS data can be affected by environmental factors, resulting in the presence of numerous outliers. Therefore, it is necessary to clean the AIS data as follows:

(1) Violation data, such as the MMSI does not conform to the 9-bit record format, latitude and longitude values fall outside the valid range, or the draft value is zero.

(2) Abnormal data, such as significant changes or lack of changes in a short or long time interval.

(3) Overlapping data, such as two adjacent trajectory points have identical values, except for the timestamp.

Furthermore, to ensure consistent time intervals between adjacent points, it is necessary to interpolate the trajectory. Based on the standard (ITU, 2010), the transmission frequency of AIS data varies depending on the navigation status of the ship, ranging from 3 seconds to 3 minutes. The basic linear interpolation method is used in this study:

$$P(t) = P(t_0) + \frac{\big(P(t_1) - P(t_0)\big)(t - t_0)}{t_1 - t_0} \tag{1}$$

where $P(t_0)$ and $P(t_1)$ represent the trajectory points at time t_0 and t_1, respectively; $P(t)$ is the interpolated point at time t, which is between t_0 and t_1.

3.3 Trajectory clustering

The clustering of trajectories can uncover the behavioral patterns of ships within a heavy traffic area. It facilitates the construction of a trajectory prediction model by grouping trajectories with similar behavioral characteristics into the same cluster. There are plenty of outliers and missing data in raw AIS data, making clustering challenging. Density-based clustering algorithms are capable of clustering irregular samples and are not sensitive to noise data. Therefore, the OPTICS clustering algorithm, one of the density clustering algorithms, is used for trajectory clustering in this paper.

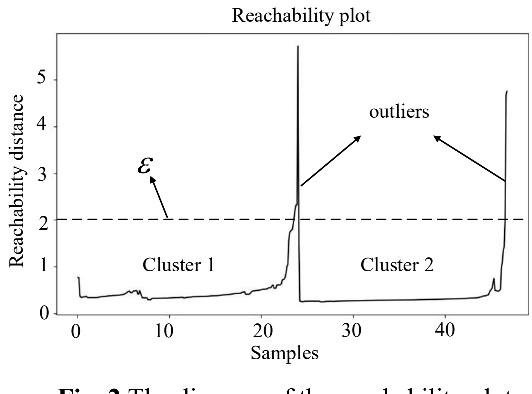


Fig. 2 The diagram of the reachability plot

The OPTICS algorithm does not explicitly generate clustering results. Instead, it performs a ranking of objects within the dataset and generates a reachability distance distribution plot known as the reachability plot. The reachability plot provides insights into the density-based structure of the data, allowing for more informed decisions in selecting suitable parameters for the clustering. As shown in Fig. 2, The horizontal axis represents the index of samples, and the vertical axis represents the reachability distance. By selecting different threshold values of reachability distance, the samples can be divided into different clusters. For example, when the threshold is set as the dashed line in Fig. 2, samples below the dashed line are classified into two clusters, while samples above the line are considered outliers.

The core objects $MinPts$ and reachability distance ε are parameters of the OPTICS algorithm. As shown in Fig. 3, $MinPts$ means the number of samples inside the circle with a radius of ε. The core idea of OPTICS is to determine the minimum radius ε, named the reachability distance, that can include a given number of samples $MinPts$. As shown in Fig. 3, the reachability distance is 2 when the $MinPts$ is set to 6, and the reachability distance is 10 when the threshold is set to 10. The expression of the reachability distance for a given $MinPts$ can be described as follows:

$$\begin{cases} N_\varepsilon(p) = \{p_i \in D | dis(p_i, p_i) \le \varepsilon\} \\ \quad |N_\varepsilon(p)| = MinPts \end{cases} \tag{2}$$

where p_i means one of the samples D; $|N_\varepsilon(p)|$ means the number of p that satisfy the upper equation; ε is the reachability distance of sample p when the core objects is $MinPts$.

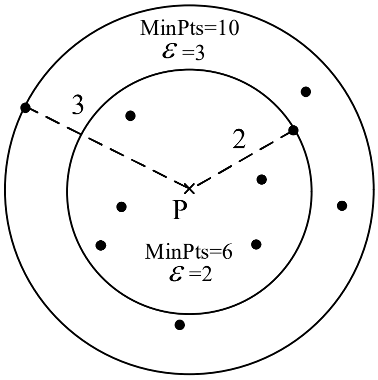


Fig. 3 The diagram of the OPTICS algorithm

In this part, the trajectory similarity is measured using the Hausdorff distance which is a distance between two sets. The equation is as follows:

$$dHau(T_a, T_b) = max\left\{\sup_{a\epsilon M}\left[\inf_{b\epsilon N}\big(dis(ab)\big)\right], \sup_{b\epsilon N}\left[\inf_{a\epsilon M}\big(dis(ba)\big)\right]\right\} \tag{3}$$

where $\sup_{a\epsilon M}\left[\inf_{b\epsilon N}\big(dis(ab)\big)\right]$ is one-way distance set from trajectory $T_a$to trajectoryT_b, $\sup_{b\epsilon N}\left[\inf_{a\epsilon M}\big(dis(ba)\big)\right]$ is one-way distance set from trajectory T_b to T_a, $dis(ab)$ and $dis(ba)$ are the Euclidean distances between record positions in trajectory T_a and T_b.

3.4 Trajectory prediction

The purpose of this part is to predict the trajectories of other vessels and provide a reference for collision avoidance decisions of the own ship. The GMM refers to a linear combination of multiple Gaussian distribution functions, which is commonly used to address the situation where data within the same dataset contains multiple different distributions. Ship trajectories consist of multidimensional data such as longitude, latitude, course, and speed. In this study, we primarily focus on the longitude and latitude of the ships and utilize the GMM to predict their trajectories.

Firstly, each clustered trajectory cluster is divided into multiple trajectory sets based on the average speed of the points on the trajectory, such as low speed (under 5 knots), moderate speed (between 5 and 10 knots), and high speed (over 10 knots). Then, linear interpolation is performed on the trajectories using the time parameter. In addition, the trajectory's latitude and longitude coordinates are transformed into a plane rectangular coordinate system through Mercator projection:

$$\begin{cases} x = lon \cdot \frac{C}{2\pi} \\ y = log\left(tan\left(\frac{(90 + lat) \cdot \pi}{360}\right)\right) \cdot \frac{C}{2\pi} \end{cases} \tag{4}$$

where lon and lat are the longitude and latitude coordinates of the trajectory point; The value of C is 20037508.34, which is the circumference of the earth.

Subsequently, a two-dimensional GMM model is established to predict the trajectories of other vessels. The joint probability distribution of a two-dimensional random variable $x = (X, Y)$ at time t can be determined by the mean vector μ and covariance matrix Σ. The mean vector represents the central location of the distribution, while the covariance matrix represents the dispersion and correlation in different directions. Specifically, the joint probability distribution is given by:

$$P(X, Y) = \frac{C}{2\pi|\Sigma|} exp\left(-\frac{1}{2}(x - \mu)\right)^T \Sigma^{-1}(x - \mu) \tag{5}$$

where $|\Sigma|$ denotes the determinant of the covariance matrix ; $(x - \mu)$ denotes the transpose of the difference vector.

The mean vector can be obtained by calculating the mean value for each dimension:

$$\mu = \begin{vmatrix} \mu_x \\ \mu_y \end{vmatrix} = \begin{vmatrix} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xf(x, y)\mathrm{d}x\mathrm{d}y \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} yf(x, y)\mathrm{d}x\mathrm{d}y \end{vmatrix} \tag{6}$$

The equation for the covariance matrix is as follows:

$$\Sigma = \begin{pmatrix} s_x^2 & s_{xy} \\ s_{xy} & s_y^2 \end{pmatrix} \tag{7}$$

Where s_x^2 and s_y^2 are the sample standard deviations of the two variables, and s_{xy} is the sample covariance:

$$s_{xy} = \frac{1}{n - 1} \sum_{i=1}^{n} (x_i - \mu_x)(y_i - \mu_y) \tag{8}$$

Based on the interpolated historical AIS trajectories, the mean μ and covariance Σ of the two-dimensional GMM can be computed for each time step, resulting in the prior distribution of the predicted trajectory. The predicted trajectory is divided into observed and predicted points. The posterior probability distribution $P(x_i, y_i)$ of the observed points can be obtained using Equation (5), and the probability distribution of predicted points also can obtained as follows:

$$P(\overline{x}, \overline{y}) = \frac{1}{N} \sum_{i=1}^{m} P(x_i, y_i) \tag{9}$$

As shown in Fig. 4, the prior distribution (μ_t, Σ_t) corresponding to the time from t_1 to t_n can be obtained by analyzing historical AIS data using Equations (6) and (7). The ellipses in the figure represent the prior distribution at each moment. By observing the trajectory points (x_t, y_t) of the target ship from t_1 to t_m, the posterior probability $P(x_t, y_t)$ corresponding to these trajectory points can be obtained. The average of these posterior probabilities, denoted as $P(\overline{x}, \overline{y})$, can be considered as the predicted probability for the future trajectory of the target ship. The predicted probability $P(\overline{x}, \overline{y})$ is represented as a trajectory band, which includes one center line and two border lines. The trajectory band represents the predicted trajectory of the target ship.

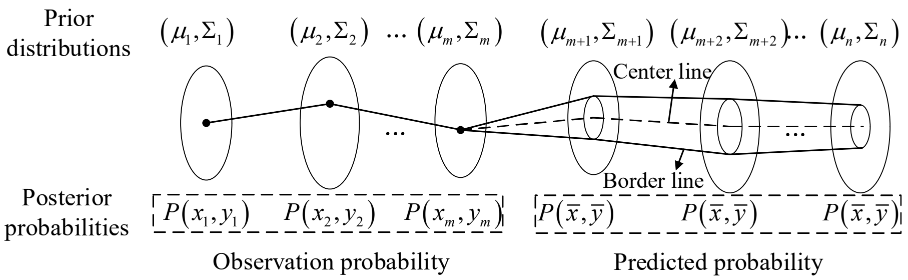


Fig. 4 The diagram of the trajectory predicted model

3.5 Collision avoidance

This part consists of two main components. The first component is the establishment of the avoidance objective function, which aims to obtain feasible avoidance action. The second component is the establishment of the ship's motion and control model, which considers the maneuverability of the ship to obtain the avoidance path corresponding to feasible avoidance action.

It is necessary to determine whether there is a collision risk within time T between the planned trajectory of the own ship and the predicted trajectory of the target ship. As shown in Fig. 5, It is considered to have collision risk between the own ship and the target ship when the trajectory point $P_i(t)$ is located within the predicted interval of point $P_j(t)$ at the same time t. The equation of collision risk is as follows:

$$P_i(t) \in P_j(t) \oplus P(\overline{x}, \overline{y}) \tag{10}$$

The avoidance action can be determined based on the position of the own ship within the predicted interval of the target ship if there is collision risk. As shown in the right subplot of Fig. 5, the own ship should perform a right turn to avoid the collision when the point of the own ship is on the right side of the predicted interval. Conversely, if the point is on the left side, the own ship should perform a left turn for avoidance. After taking an avoidance action, the nearest planned waypoint can be considered as the target place, guiding the ship back to the planned route.

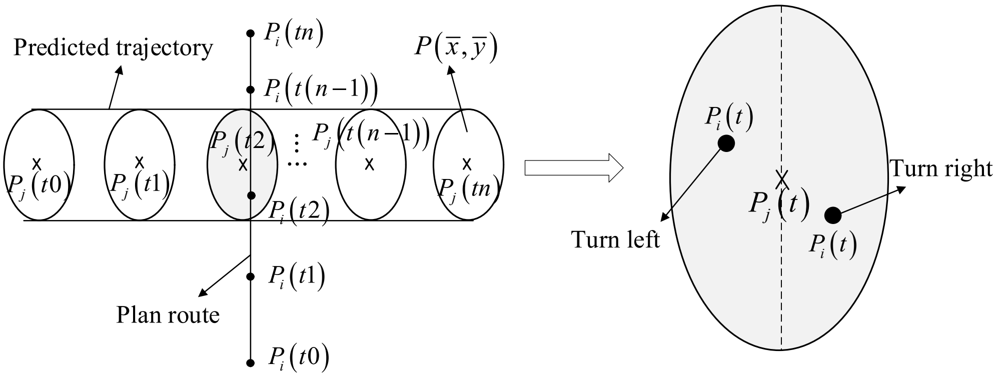


Fig. 5 The diagram of collision avoidance

The maneuvering motion of the ship exhibits time-varying and nonlinear characteristics. It typically relies on reliable mathematical models to predict the ship's motion. When it comes to collision avoidance, only the primary motions that affect the avoidance, namely yaw, surge, and sway, are considered. Therefore, this paper adopts the MMG motion model which only considered the above factors:

$$\begin{cases} (m+m_x)\dot{\mu} - (m+m_y)vr = X_H + X_p + X_R \\ (m+m_y)\dot{v} - (m+m_x)\mu r = Y_H + Y_p + Y_R \\ (I_{ZZ} + J_{ZZ})\dot{r} = N_H + N_p + N_R \end{cases} \tag{11}$$

where v and μ are the lateral and surge velocity, respectively; m is ship mass, and m_x, and m_y are the added masses of the x and y-axis direction; r is yaw rate; I_{ZZ} and J_{ZZ} are the moments of the ship's inertia around z –axes and the added moment of inertia around the z-axis, respectively; X_H, Y_H, N_H are ship hydrodynamic moments and forces; X_p, Y_p, N_p are propeller moments and forces; X_R, Y_R, N_R are rudder forces and moments. Zhao et al. [42] provide a detailed introduction to the parameters of the "YuKun". This paper establishes an MMG motion model based on these parameters.

The consideration of the course control system is also important. In this paper, the classical PID navigation control system, which is commonly installed on most sea-going ships, is adopted:

$$\mu(t) = K_p e(t) + K_i \int_0^t e(t)\mathrm{d}t + K_d \frac{\mathrm{d}e(t)}{\mathrm{d}t} \tag{12}$$

where K_p, K_i, and K_d represent the proportional gain, integral time, and derivative time, respectively. The parameters set in this paper are $K_p = 1.47,\ K_i = 0.015, K_d = 95.2$, as referenced in Zhao et al. [42].

4. Case study

4.1 Case design

To demonstrate the effectiveness of the proposed method, one case study is illustrated in this section. The experimental area, with longitude ranging from 122.4895° to 122.588° and latitude ranging from 30.82798° to 30.92245°, is illustrated in Fig. 6. The area consists of two main routes, namely the north-south route and the east-west route, with intersections between them. Assuming that the own ship is sailing westward on the east-west route. Meanwhile, there is a target ship sailing northward on the north-south route. In this situation, the own ship needs to predict the trajectory of the target ship, assess the risk of collision, and take avoidance action if there is a collision risk. In this experiment, the collision avoidance rules have not been considered. thus, it is assumed that the target ship will maintain its course, even if it is the give-way ship.

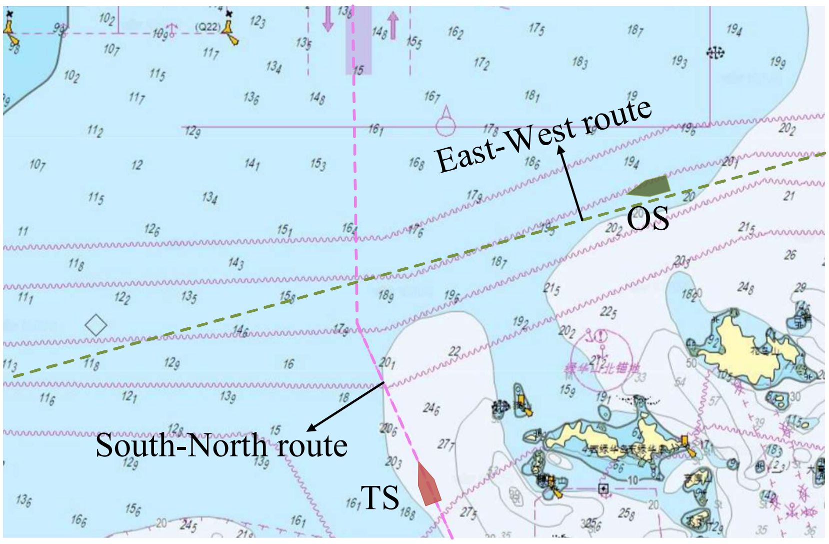
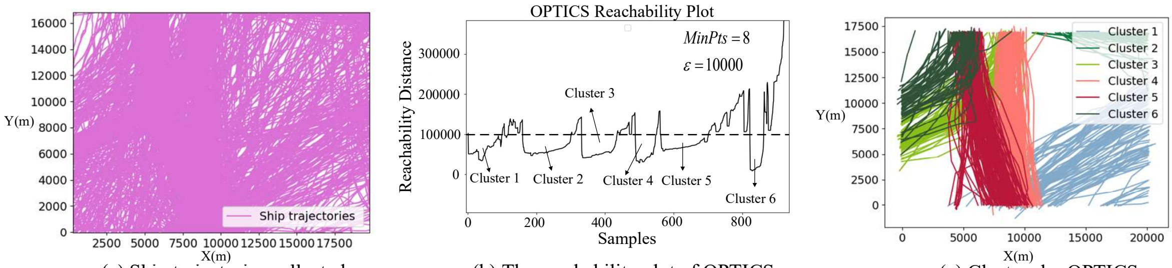


Fig. 6 The experimental scenario of the case study

To understand the ship traffic characteristics of the water area, this paper collected AIS data from the Zhoushan Maritime Safety Administration of China for the week of July 21 to July 27, 2021. First, the raw data will be decoded, cleaned, and interpolated according to the method proposed in this paper. Second, the preprocessed trajectories will be clustered using the OPTICS algorithm, matching the target ship with the corresponding trajectory cluster. Then, a GMM will be used to predict the trajectory of the target ship. Finally, collision risk will be analyzed based on ship maneuverability, and effective evasive actions will be taken.

4.2 Result of case

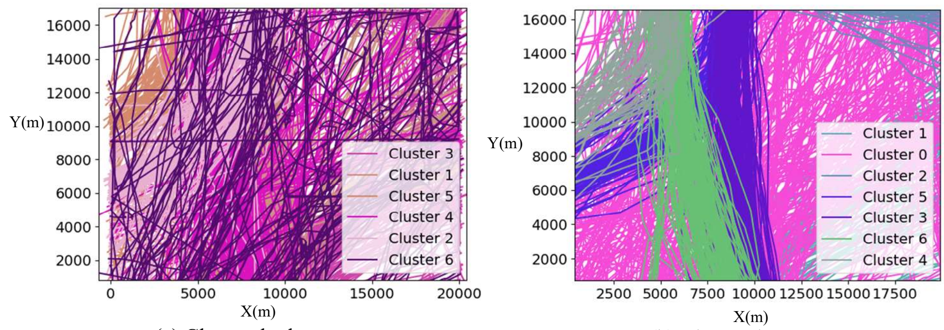
There are a total of 929 ship trajectories in the experimental area after decoding and preprocessing the collected data, as shown in Fig. 7(a). The study applies the OPTICS algorithm to cluster the trajectories with the parameters of $MinPts = 8$ and $\varepsilon = 10000$, indicating that the number of trajectories is 8 and the Euclidean distance is 10000. The value of $MinPts$ influences the distribution of the curve, while $\varepsilon = 10000$ is primarily used for dividing trajectory clusters. The corresponding reachability distance plot, where the dashed line represents $\varepsilon = 10000$, is shown in Fig. 7(b). The curve ascending above the dashed line is classified as noise, while the curve descending below the dashed line is divided into six distinct sections, corresponding to the six trajectory clusters in the Fig. 7(c). The cluster 4 with orange and cluster 5 with red have the highest match with the north-south routes, corresponding to the northbound and southbound routes, respectively. The other clusters are customary routes of the east-west route.



(a) Ship trajectories collected (b) The reachability plot of OPTICS (c) Clusters by OPTICS

Fig. 7 The clustering result of case study

Fig. 8 shows the clustering results using two other common algorithms, K-means and DBSCAN. It can be observed from Fig. 8(a) that the clustering effect of K-means is not satisfactory. This is because K-means are easily affected by noise, and ship trajectory data is relatively chaotic, making K-means unsuitable for clustering of ship trajectories. Fig. 8(b) shows the results of clustering using the DBSCAN algorithm. It is evident from the figure that the clustering effect is superior to that of K-means. However, DBSCAN cannot eliminate ineffective clustering trajectories. Compared to these two algorithms, OPTICS has the following two advantages: 1) Clustering parameters $MinPts$ and ε can be accurately identified by the OPTICS reachability plot. 2) It is insensitive to noise and can eliminate ineffective clustering trajectory clusters.



(a) Clusters by k-means (b) Clusters by DBSCAN

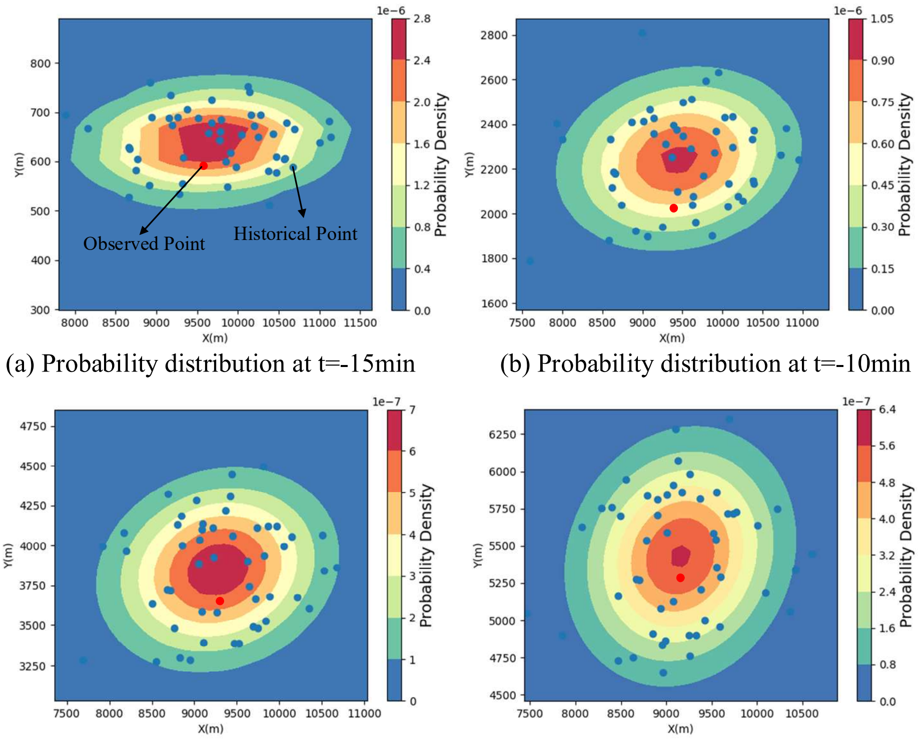
Fig. 8 The clustering results by other methods

In this case, the cluster 4 is chosen as the foundational dataset for trajectory prediction since the target ship is heading north. A linear interpolation is performed with a time interval of t=60s to obtain the subsequent trajectory points $\sum_{t=1}^{n} P(t)$. Then, the mean μ_t and covariance matrix Σ_t of all trajectory points at the time t are calculated separately. The probability distribution $\sum_{t=1}^{n}(\mu_t, \Sigma_t)$ of trajectory points $\sum_{t=1}^{n} P(t)$ are used as prior information for the trajectory prediction model. As shown in Table 1, this is the prior distribution (μ_t, Σ_t) corresponding to each time step after interpolation.

As shown in Fig. 9, the posterior probabilities of the target ship's trajectory points at t=-15 min, -10 min, -5 min, and 0 min before the current time of the own ship are illustrated. In the figures, the blue dots and red dots represent the historical trajectory points and observed points, respectively. The probability of red dots is the posterior probabilities of these times. As shown in Fig. 10, the posterior probabilities of all observation points are obtained, and the average value of 5.60e-07 is taken as the predicted probability for the trajectory prediction model.

Table 1 The results of prior distributions

Time (min)	Mean (μ_t)	Covariance matrix (Σ_t)
10	$\begin{vmatrix} 9338 \\ 3211 \end{vmatrix}$	$\begin{vmatrix} 5.8553e{+}05 & 3.2217e{+}04 \\ 3.2217e{+}04 & 4.5712e{+}05 \end{vmatrix}$
20	$\begin{vmatrix} 9091 \\ 6377 \end{vmatrix}$	$\begin{vmatrix} 3.7280e{+}05 & 3.9464e{+}04 \\ 3.9464e{+}04 & 2.7681e{+}05 \end{vmatrix}$
30	$\begin{vmatrix} 8955 \\ 9503 \end{vmatrix}$	$\begin{vmatrix} 3.5882e{+}05 & 1.9132e{+}04 \\ 1.9132e{+}04 & 8.0296e{+}05 \end{vmatrix}$
40	$\begin{vmatrix} 8870 \\ 12609 \end{vmatrix}$	$\begin{vmatrix} 3.3445e{+}05 & 9.5729e{+}04 \\ 9.5729e{+}04 & 1.7464e{+}06 \end{vmatrix}$



(a) Probability distribution at t=-15min (b) Probability distribution at t=-10min

(c) Probability distribution at t=-5min (d) Probability distribution at t=0min

Fig. 9 The example of prior and posterior probabilities

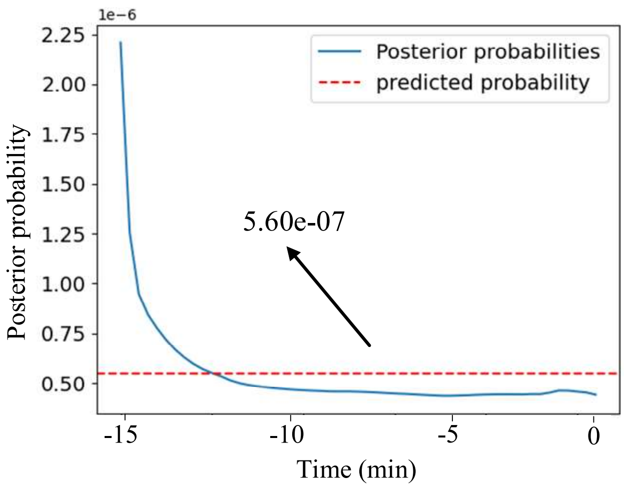
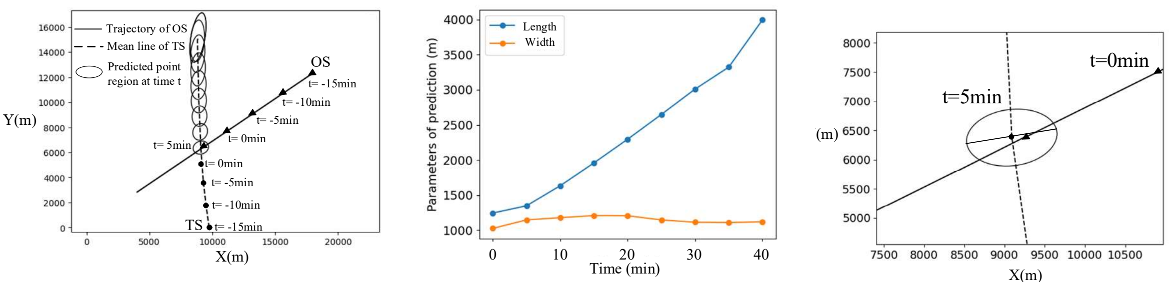


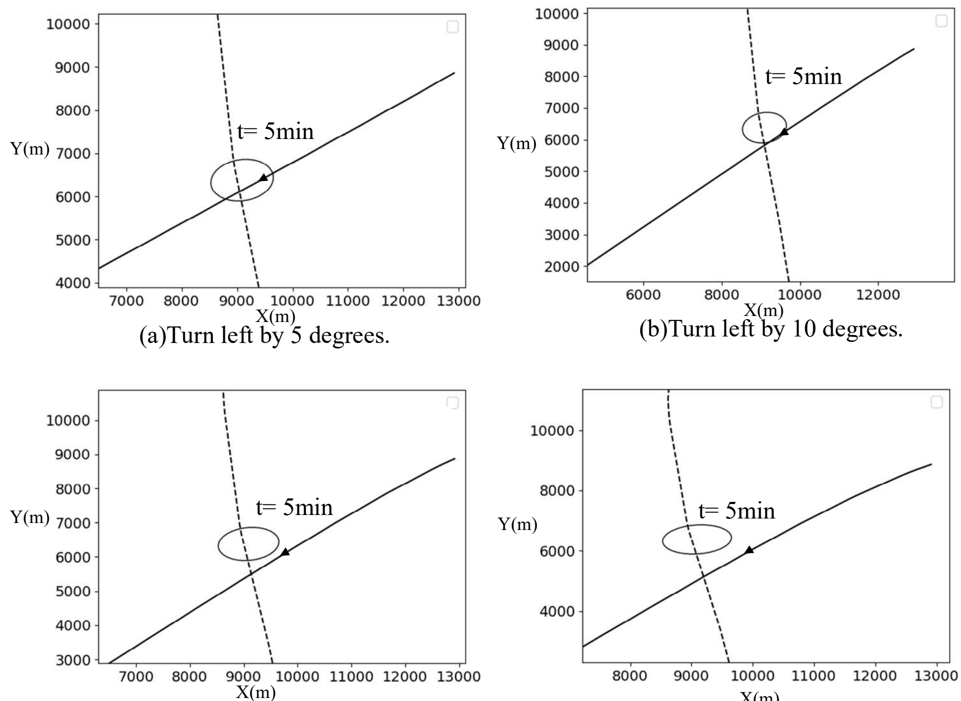
Fig.10 The results of posterior probabilities

Fig. 11(a) shows the predicted trajectory results of the target ship. The black ellipses in the figure represent the predicted trajectory point regions at different times. The dashed lines connect the means of the point regions. The black line represents own ship's trajectory while maintaining its current movement state. This paper predicts the trajectory of the target ship by fitting it with an ellipse. The predicted results of the ellipse's major (length) and minor (width) axes are shown in Fig. 11(b). It is evident that the length of the predicted ellipse increases over time, which means that the accuracy of the prediction decreases as time progresses. When the time reached 10 minutes and 30 minutes, the errors of the predicted major axis (length) of the ellipse exceeded 1500m and 3000m, respectively. As shown in Fig. 11(c), the mean line of TS and the trajectory of OS intersects at t=5 minutes, and the intersection point falls exactly within the predicted trajectory point region of TS at this time. Therefore, the own ship needs to take evasive action to avoid entering the predicted trajectory area. From the perspective of the own ship, since the intersection point is on the left side of the predicted region, own ship should perform a left turn to avoid according to the avoidance method mentioned in this paper.



(a) The prediction trajectory of TS (b) The parameters of prediction (c) The prediction result at t=5min

Fig.11 The results of the predicted trajectory



(a)Turn left by 5 degrees. (b)Turn left by 10 degrees.

(c)Turn left by 15 degrees. (d)Turn left by 20 degrees.

Fig.12 The results of collision avoidance

The paper utilizes the MMG and PID models to simulate ship maneuverability, with increments of 5 degrees. The simulation results are presented in Fig. 12. It's evident that when making left turns of 5° and 10°, the intersection point of the own ship's trajectory and the mean line inside the predicted area of the target ship at t=5 min, rendering these avoidance maneuvers ineffective. Conversely, when the turning angle reaches

15° and 20°, the intersection point falls outside the predicted area of the target ship. Hence, the own ship should adopt a left turn exceeding 15° to mitigate the risk of collision with the target ship.

5. Conclusions

In this paper, a collision avoidance decision-making research in coastal waters considering the uncertainty of target ships was proposed. The contribution of this paper is two-fold:1) This paper proposed a complete approach to extract trajectory clusters of both own and target ships from raw AIS data ship; 2) The collision avoidance decision-making approach that takes into account the trajectory uncertainty of the target ship proposed in this paper. The model design of this paper mainly includes data preprocessing, trajectory clustering, trajectory prediction, and collision avoidance. Accurate ship trajectories are obtained by preprocessing the raw data. Then, the processed trajectories are clustered using the OPTICS algorithm and Hausdorff distance, enabling the matching of the target ship with the corresponding trajectory cluster for trajectory prediction. Subsequently, based on the matched trajectory cluster, a mixed Gaussian model is utilized to calculate the prior probability distribution of the prediction model. The posterior probability of the observed target trajectory determines the average probability of the prediction model, thus establishing a trajectory prediction model that considers the uncertainty of the target ship. Finally, ship maneuverability is simulated using the MMG and PID models, and a collision avoidance decision-making model for ships is constructed.

To showcase the efficacy of the proposed method, a case study using one-week AIS data was presented. The experimental outcomes indicate that the approach effectively predicts the trajectory of the target ship and facilitates informed collision avoidance decision-making. However, this method also has the following shortcomings: 1) It does not consider the collision avoidance rules; 2) The accuracy of target ship trajectory prediction needs to be improved, with the prediction accuracy decreasing as the prediction time increases. 3) The research is only applicable to ship-ship and is not yet capable of avoiding collisions among multiple ships.

For future investigations concerning collision avoidance in coastal waters, the following aspects are considered. First, enriching the collision avoidance decision-making model by integrating collision avoidance rules and navigation practice. Secondly, enhancing the precision of the prediction model to obtain more accurate trajectory predictions. Third, the applicability of the extended method in collision avoidance for multiple ships. Finally, the avoidance timing should be considered in future research.

Acknowledgment

The work presented in this study is financially supported by the Natural Science Foundation of Fujian Province of China (No. 2023J01326) and the Fujian Provincial Department of Education (JAT190296).

REFERENCES

[1] Wu, G., Li, Y., Jiang, C., Wang, C., Guo, J., Cheng, R., 2022. Multi-vessels collision avoidance strength for autonomous surface vehicles based on genetic algorithm in congested port environment. *Brodogradnja*, 73(3), 69-91. https://doi.org/10.21278/brod73305

[2] EMSA, Preliminary annual overview OF marine casualties and incidents 2014-2019, 2020. http://www.emsa.europa.eu/ accident-investigation-publications/annual-overview.html

[3] Zhu, Z., Lyu, H., Zhang, J., Yin, Y., 2021. An Efficient Ship Automatic Collision Avoidance Method Based on Modified Artificial Potential Field. *Journal of Marine Science and Engineering*, 10(1), 3. https://doi.org/10.3390/jmse10010003

[4] Zhang, K., Huang, L., He, Y., Wang, B., Chen, J., Tian, Y., Zhao, X., 2023. A real-time multi-ship collision avoidance decision-making system for autonomous ships considering ship motion uncertainty. *Ocean Engineering*, 278, 114205. https://doi.org/10.1016/j.oceaneng.2023.114205

[5] Sano ,M., 2023. Mathematical model and simulation of cooperative manoeuvres among a ship and tugboats. *Brodogradnja*, 74(2), 127-148. https://doi.org/10.21278/brod74207

[6] Zhang, S., Shi, G., Liu, Z., Zhao, Z., Wu, Z., 2018. Data-driven based automatic maritime routing from massive AIS trajectories in the face of disparity. *Ocean Engineering*, 155, 240-250. https://doi.org/10.1016/j.oceaneng.2018.02.060

[7] Zhao, L. Shi, G., 2019. A novel similarity measure for clustering vessel trajectories based on dynamic time warping. *The Journal of Navigation*, 72(2), 290-306. https://doi.org/10.1017/S0373463318000723

[8] Zhen, R., Jin, Y., Hu, Q., Shao, Z., Nikitakos, N., 2017. Maritime anomaly detection within coastal waters based on vessel trajectory clustering and Naïve Bayes Classifier. *The Journal of Navigation*, 70(3), 648-670. https://doi.org/10.1017/S0373463316000850

[9] Bai, Q., Jin, C., 2022. A K-Means and Ant Colony Optimization-Based Routing in Underwater Sensor Networks. *Mobile Information Systems*, 2022(12), 4465339. https://doi.org/10.1155/2022/4465339

[10] Pallotta, G., Vespe, M., Bryan, K., 2013. Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction. *Entropy*, 15(6), 2218-2245. https://doi.org/10.3390/e15062218

[11] Yang, J., Liu, Y., Ma, L., Ji, C., 2022. Maritime traffic flow clustering analysis by density based trajectory clustering with noise. *Ocean Engineering,* 249, 111001. https://doi.org/10.1016/j.oceaneng.2022.111001

[12] Zhao, L., Shi, G., 2019. A trajectory clustering method based on Douglas-Peucker compression and density for marine traffic pattern recognition. *Ocean Engineering*, 172, 456-467. https://doi.org/10.1016/j.oceaneng.2018.12.019

[13] Shen, H., Tang, H., Yin, Y., 2022. A novel method for ship trajectory clustering. *International Journal of Naval Architecture and Ocean Engineering,* 14, 100474. https://doi.org/10.1016/j.ijnaoe.2022.100474

[14] Cao, J., Liang, M., Li, Y., 2018. PCA-based hierarchical clustering of AIS trajectories with automatic extraction of clusters, *2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA).* 9-12 March, Shanghai, China, 448-452. https://doi.org/10.1109/ICBDA.2018.8367725

[15] Vicente-Cera, I., Acevedo-Merino, A., Nebot, E., 2020. Analyzing cruise ship itineraries patterns and vessels diversity in ports of the European maritime region: A hierarchical clustering approach. *Journal of Transport Geography*, 85, 102731. https://doi.org/10.1016/j.jtrangeo.2020.102731

[16] Wang S, He Z., 2021, A Fast Clustering Algorithm for Ship Trajectories Based on Layered Modeling *Journal of Wuhan University of Technology: Transportation Science and Engineering Edition*, 45(3), 1-6.

[17] Liang M., 2023. Research on Multi Scenario Ship Trajectory Prediction Method Driven by Knowledge and Data. Wuhan University of Technology, 2023.

[18] Zhao, L., Shi, G., Yang, J., 2017. An adaptive hierarchical clustering method for ship trajectory data based on DBSCAN algorithm, 2*017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)*. 10-12 March, Beijing, China, 329-336. https://doi.org/10.1109/ICBDA.2017.8078834

[19] Ma, Q., Zhou, Y., Liu, L., 2022. Review and comparison of the demand analysis methods of maritime emergency resources. *Brodogradnja,* 73(1), 141-162. https://doi.org/10.21278/brod73108

[20] Xiao, Z., Fu, X., Zhang, L., Zhang, W., Goh, G., 2022. Big Data Driven Vessel Trajectory and Navigating State Prediction With Adaptive Learning, Motion Modeling and Particle Filtering Techniques. *IEEE Transactions on Intelligent Transportation Systems*, 23(4), 3696-3709. https://doi.org/10.1109/TITS.2020.3040268

[21] Perera, L., Soares, C., 2010. Ocean vessel trajectory estimation and prediction based on extended Kalman filter. *The Second International Conference on Adaptive and Self-Adaptive Systems and Applications.* 21-26 November, Lisbon, Portugal, 14-20.

[22] Virjonen, P., Nevalainen, P., Pahikkala, T., 2018. Ship movement prediction using k-NN method. *2018 Baltic Geodetic Congress (BGC Geomatics)*. 21-23 June, Olsztyn, Poland, 304-309. https://doi.org/10.1109/BGC-Geomatics.2018.00064

[23] Liu, J., Shi, G., Zhu, K., 2019. Vessel trajectory prediction model based on AIS sensor data and adaptive chaos differential evolution support vector regression (ACDE-SVR). *Applied Sciences*, 9(15), 2983. https://doi.org/10.3390/app9152983

[24] Rong, H., Teixeira, A., Soares, C., 2019. Ship trajectory uncertainty prediction based on a Gaussian Process model. *Ocean Engineering*, 182, 499-511. https://doi.org/10.1016/j.oceaneng.2019.04.024

[25] Lazarowska, A., 2020. Comparison of discrete artificial potential field algorithm and wave-front algorithm for autonomous ship trajectory planning. *IEEE Access*, 8, 221013-221026. https://doi.org/10.1109/ACCESS.2020.3043539

[26] Silveira, P., Teixeira, A., Guedes, C., 2019. AIS based shipping routes using the dijkstra algorithm. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, 13(3), 176-189. https://doi.org/10.12716/1001.13.03.11

[27] Karbowska, C., Koszelew, J., Ostrowski, K., 2019. Beam search Algorithm for ship anti-collision trajectory planning. *Sensors*, 19(24), 5338. https://doi.org/10.3390/s19245338

[28] He, Y., Liu, X., Zhang, K., Mou, J., Liang, Y., Zhao, X., Wang, B., Huang, L., 2022. Dynamic adaptive intelligent navigation decision making method for multi-object situation in open water. *Ocean Engineering*, 253, 111238. https://doi.org/10.1016/j.oceaneng.2022.111238

[29] Huang, C., Li, W., Xiao, C., Liang, B., Han, S., 2018. Potential field method for persistent surveillance of multiple unmanned aerial vehicle sensors. *International journal of distributed sensor networks*, 14(1), 1550147718755069. https://doi.org/10.1177/1550147718755069

[30] Wang, S., Fang, M., Huang, C., 2019. Vertical obstacle avoidance and navigation of autonomous underwater vehicles with H∞ controller and the artificial potential field method. *The Journal of Navigation*, 72(1), 207-228. https://doi.org/10.1017/S0373463318000589

[31] Singh, Y., Sharma, S., Sutton, R., 2018. A constrained A* approach towards optimal path planning for an unmanned surface vehicle in a maritime environment containing dynamic obstacles and ocean currents. *Ocean Engineering*, 169, 187-201. https://doi.org/10.1016/j.oceaneng.2018.09.016

[32] Xue, H., 2022. A quasi-reflection based SC-PSO for ship path planning with grounding avoidance. *Ocean Engineering*, 247, 110772. https://doi.org/10.1016/j.oceaneng.2022.110772

[33] Xie, S., Chu, X., Zheng, M., Liu, C., 2019. Ship predictive collision avoidance method based on an improved beetle antennae search algorithm. *Ocean Engineering*, 192, 106542. https://doi.org/10.1016/j.oceaneng.2019.106542

[34] Zhang, K., Huang, L., Liu, X., Chen, J., Zhao, X., Huang, W., He, Y., 2022, A novel decision support methodology for autonomous collision avoidance based on deduction of manoeuvring process. *Journal of Marine Science and Engineering*, 10, 765. https://doi.org/10.3390/jmse10060765

[35] Liu, J., Zhang, J., Yan, X., Soares, C., 2022. Multi-ship collision avoidance decision-making and coordination mechanism in Mixed Navigation Scenarios. *Ocean Engineering*, 257, 111666. https://doi.org/10.1016/j.oceaneng.2022.111666

[36] Wang, T., Wu, Q., Zhang, J., Wu, B., Wang, Y., 2020. Autonomous decision-making scheme for multi-ship collision avoidance with iterative observation and inference. *Ocean Engineering*, 197, 106873. https://doi.org/10.1016/j.oceaneng.2019.106873

[37] Fiskin, R., Atik, O., Kisi, H., 2021. Fuzzy domain and meta-heuristic algorithm-based collision avoidance control for ships: Experimental validation in virtual and real environment. *Ocean Engineering*, 220, 108502. https://doi.org/10.1016/j.oceaneng.2020.108502

[38] Xie, S., Chu, X., Zheng, M., Liu, C., 2019. Ship predictive collision avoidance method based on an improved beetle antennae search algorithm. *Ocean Engineering*, 192, 106542. https://doi.org/10.1016/j.oceaneng.2019.106542

[39] Lazarowska, A., 2020. A discrete planning approach in collision avoidance for smart ships. *Procedia Computer Science*, 176, 380-389. https://doi.org/10.1016/j.procs.2020.08.039

[40] Chen, C., Chen, X., Ma, F., Zheng, X., Wang, J., 2019. A knowledge-free path planning approach for smart ships based on reinforcement learning. *Ocean Engineering*, 189, 106299. https://doi.org/10.1016/j.oceaneng.2019.106299

[41] Gao, X., Dong, Y., 2023. An Optimized Path Planning Method for Container Ships in Bohai Bay Based on Improved Deep Q-Learning. *IEEE Access,* 11, 91275-91292. https://doi.org/10.1109/ACCESS.2023.3307480

[42] Zhao B, Zhang X, Liang C., 2021. An improved model predictive control for path-following of USV based on global course constraint and event-triggered mechanism. *IEEE Access*, 9, 79725-79734. https://doi.org/10.1109/ACCESS.2021.3084844